心理科学进展 2020, Vol. 28, No. 8, 1219–1231
Advances in Psychological Science DOI: 10.3724/SP.J.1042.2020.01219

·研究构想(Conceptual Framework)·

动、静态视觉信息在真实世界视觉搜索中的作用*

潘 静 张慧远 陈东濠 徐宏格

(中山大学心理学系, 广州 510006)

摘 要 真实环境中的视觉搜索是人和动物赖以生存的重要能力。目前的视觉搜索研究多使用静态的观察者和静止的二维搜索对象, 侧重于探究注意在搜索中的作用; 现有的视觉搜索理论模型主要概括了影响搜索的自上而下的注意因素, 而将自下而上影响因素简单归结为影像显著性, 然而在真实环境中, 观察者或搜索对象是可以运动的, 搜索时可利用的视觉信息包括动态光流和静态影像结构信息。已有的视觉识别研究发现这两种信息相结合可以使观察者准确持久地识别场景、事件和三维结构。在现有视觉搜索理论模型中引入两种视觉信息可以较好还原真实环境中的搜索任务。我们提出研究构想和实验方案, 探究利用动、静态视觉信息的视觉搜索过程, 从而完善现有的视觉搜索模型。我们认为充分利用环境信息可以提高搜索效率, 且在视觉搜索训练和智能搜索设计等方面有重要的应用价值。

关键词 视觉搜索; 光流; 生物运动; 生态知觉理论
分类号 B842

1 引言

1.1 研究背景

视觉搜索(visual search)指通过观察、利用视觉信息从众多事物中找到某一目标。这是人和动物赖以生存的关键能力之一。对于视觉搜索过程及其认知机制的研究有助于人们发现影响搜索效率的因素, 从而有针对性地制定视觉搜索训练计划, 提高搜索准确性和搜索速度。在经典的视觉搜索实验中, 一个搜索目标和几个干扰子随机且独立地呈现在电脑屏幕上(Treisman & Gelade, 1980; Koch & Ullman, 1987; Duncan & Humphreys, 1989; Wolfe & Gancarz, 1997), 目标和干扰子之间在一个或几个维度上有所区分, 如在多个字母S中找出字符$, 或者在众多不同颜色的几何形状中找出红色圆圈。与真实世界视觉搜索任务相比, 这类实验室任务中的搜索目标、干扰子和背景信息都过于简单。近年来, 研究者通过使用真实物体的静态图片作为搜索对象(或呈现在单一颜色背景中, 或嵌于静态场景中)增加视觉搜索实验范式的生态效度。这些研究发现场景结构/要义(Torralba, Oliva, Castelhano, & Henderson, 2006; Henderson & Hayes, 2017)、搜索历史(Ort, Fahrenfort, & Olivers, 2017; Wolfe, Cain, & Aizenman, 2019)、目标频率(Wolfe & Wert, 2010; Wolfe, Boettcher, Josephs, Cunningham, & Drew, 2015) 以及目标价值(Hickey, Chelazzi, Theeuwes, & Geng, 2014; Ehinger & Wolfe, 2016)等因素都会影响视觉搜索。

然而, 真实世界里的视觉搜索远比实验室任务复杂。具体来说, 传统视觉搜索实验范式中的观察者、搜索对象和搜索环境多是静止的, 现实中它们是可以移动的。其次, 实验室任务中通常使用简单抽象的二维图片作为目标物和干扰子。而现实中的搜索对象是三维的, 它们更加复杂、可能有数不尽的特征组合。更重要的是, 真实场景中搜索的物体外观会随着观察距离、观察角度、光影、物体或观察者的运动而改变, 甚至会被遮挡掩蔽(Foulsham & Underwood, 2009)。即便近年来有些研究使用真实照片或虚拟现实场景作为搜

收稿日期: 2020-02-11
* 国家自然科学基金(31970988)、广东省基础与应用基础研究基金(2020A1515010630)、中山大学青年教师重点培养项目(19wkzd22)。
通信作者: 潘静, E-mail: panj27@mail.sysu.edu.cn

索环境，研究者依然将丰富的三维动、静态信息简化成物体的二维投影，以简单影像作为搜索过程的起点，致使至少一半的视觉信息(动态信息)没有被纳入现有的视觉搜索模型。

因此，很多研究者对实验室任务的效度提出质疑，认为基于实验室任务得出的结论无法完全解释真实世界中的行为表现(Broadbent, 1991; Kingstone, Smilek, & Eastwood, 2008)，甚至可能起到反作用，将研究带入错误的方向(Kingstone, Smilek, Ristic, Friesen, & Eastwood, 2003)。Kingston和同事(Kingston et al, 2003)明确提出视觉搜索研究必须开拓新方向，必须将观察者、观察对象、搜索任务和环境结合，尤其是要考虑观察者的状态和自然环境特征，从具身的角度研究视觉搜索。这与Gibson生态知觉理论的核心思想(Gibson, 1958; Gibson, 1979/1986)不谋而合。

视觉搜索过程涉及到知觉(Treisman, 1982; Theeuwes, Kramer, & Belopolsky, 2004)、注意(Kristjánsson, Jóhannesson, & Thornton, 2014; Wolfe & Horowitz, 2017)、工作记忆(Drew, Boettcher, & Wolfe, 2016; Drew, Boettcher, & Wolfe, 2017)、长时记忆(Woodman & Chun, 2006; Võ & Wolfe, 2015)等认知活动。这些认知活动并非相互独立地存在于视觉搜索的过程中，而是互相交织渗透，形成了一个延续的过程。然而，大部分视觉搜索的研究中，对搜索过程和机制的讨论都侧重注意的引导、分配或捕捉。相较之下，虽然知觉贯穿了搜索过程中早期的特征注册阶段(Treisman, Sykes, & Gelade, 1977; Treisman & Gormican, 1988; Wolfe, & Gray, 2007)、中期的目标噪声分离阶段(Eriksen & Schultz, 1979)和后期的序列识别阶段(Treisman & Gormican, 1988; Wolfe, Cave, & Franzel, 1989; Wolfe, & Gray, 2007)，对搜索起到重要作用，但对搜索过程中知觉信息处理的研究却非常少。Nakayama 和 Martini (2011)提出，知觉研究尤其是物体识别可以帮助我们理解视觉搜索。他们认为物体识别和视觉搜索本质上都是模式识别(pattern recognition)。物体识别需要利用很多维度上的特征给一个物体分类；视觉搜索任务则利用少数维度上的特征区分多个物体。两种任务本质上是相同的，它们只是维度数量和物体数量的权衡(trade-off)。由此可见，识别和搜索是一个连续的两个极端，它们之间联系紧密，因此帮助物体识别的信息也可以帮助视觉搜索。基于此，本研究将从生态知觉视角出发，提出实验构想，探究动、静态视觉信息对搜索的影响，对现有的视觉搜索模型进行补充、完善。

1.2 研究意义

在理论层面，本研究打破固有思维模式，将搜索还原到最真实的场景，提出观察者和观察对象的状态会影响搜索行为，而这种影响是基于信息的：观察者和观察对象的状态不同，产生的视觉信息不同；视觉信息不同，则搜索行为不同。和其它知觉任务一样，视觉搜索是一个主动的、动态的、具身的过程，涉及搜索对象、环境背景、观察者自身等众多方面。本研究旨在完善视觉搜索理论体系，找出真实世界里静止或运动的观察者如何利用光流和影像结构信息搜索物体或事件，并通过训练提高观察者提取、整合、利用视觉信息的能力，从而提升搜索效率。

在应用方面，人生活在三维环境中，环境是不断变化的，人也是在运动中的。观察者如何在运动的状态下寻找目标，如何从一群运动的人中找出一个运动的目标，这些是重要的认知任务，在真实的、动态的环境中有广泛应用。例如，在公共安全方面，警察需要从视频监控中快速找到目标，急救人员需要在人员密集且流动的地方迅速找到需要帮助的人，这些都需要利用动、静态视觉信息进行搜索。现有搜索理论主要是基于静态影像信息，流程大概是先“记住”一个目标图像，然后从很多图像中找出目标。但是当人在环境中运动时，情况很复杂，面孔、外形等作为识别依据的图像经常会被遮挡，也很容易被修饰。仅基于影像的视觉搜索不是最有效的，容易出现漏报(miss)或者虚报(false alarm)。倘若在搜索过程中可以得到并利用影像和运动信息，可能搜索效率会更高。因为运动形态取决于物理特征(如质量、长度、摩擦系数等，每个人体型不一样，走路的姿态也不一样)，并且运动形态很难被改变。所以影像结构信息和运动产生的光流信息在视觉搜索中都很重要。在本研究中，我们验证观察者能否使用运动产生的视觉信息进行搜索，并完善基于动、静态视觉信息进行搜索的认知模型。

本研究结果的应用价值总结为：(1)做出有针对性的训练方案，着重训练对有效视觉信息的提取及整合，提高搜索准确率，降低搜索时间；(2)

将人类行为学研究结果和机器学习相结合，使自动化搜索高效低耗。现有的智能搜索算法主要基于影像信息，原则上讲影像越高清、帧数越多，搜索越准确。这样的搜索计算负荷大，对处理器、散热装置和电池的要求很高，不适用于小型便携设备。但是对光流的处理只需要低频空间信号，对清晰度没有很高要求。所以相比传统算法，基于光流的搜索具有运算速度快、计算负荷低，发热少、能耗低等优点，非常适用于便携式搜索装置或无人机搜索。本研究突破固有思维窠臼，将动态光流信息引入视觉搜索，在提高人类搜索效率、改善人工智能搜索算法等应用方面起到推进作用，有较高的应用价值和社会意义，正如 Wolfe 所说“我们的健康和安全部分寄托于成功的搜索”(“our health and safety rely, in part, on successful search”, Wolfe, 2003, p.75)。

2　国内外研究现状

2.1　视觉搜索理论及发展

Anne Treisman 以其突破性的特征整合理论(Feature Integration Theory, FIT)开启了当代视觉搜索研究的进程。FIT 将搜索过程划分为两个阶段：前注意阶段，特征首先在视野中自动和并行加工；随后的注意阶段，则借助注意来绑定特征，从而对物体进行序列识别(Treisman & Gormican, 1988)。Duncan 和 Humphreys (1989)不同意 Treisman 关于并行搜索和序列搜索之间的二分法，转而提出了相似性理论(similarity theory)。他们认为，当干扰子是同质的并且与目标非常不同时，视觉搜索任务很容易，反之则难。为了填补前注意阶段中特征对于注意分配的引导机制的空白，Wolfe 等(1989)修改了 FIT 并提出了引导搜索模型。在该模型中，对注意的引导分为自上而下和自下而上两个部分。其中，自下而上的注意引导关注的是刺激的局部对比度或物理显著性信息；自上而下的注意引导关注的是在不同特征上当前项目与目标的匹配程度。最终的搜索效率是两种引导的加权和。引导搜索模型影响巨大，此后几乎所有的视觉搜索理论研究基本都是在该模型的框架内细化、论证。

近年来，借助技术(尤其是便携式眼动、虚拟现实)和算法(贝叶斯估计、网络模型等)的发展，越来越多的视觉搜索研究通过增加背景和搜索对象的复杂程度(如使用真实物体图片或场景照片)来提高实验室任务的生态性。Wolfe 的引导搜索模型也不断地得到扩展(Wolfe 本人也不断对引导性搜索模型进行改良：Guided Search 2.0 -- Wolfe, 1994; Guided Search 3.0 -- Wolfe & Gancarz, 1997; Guided Search 4.0 -- Wolfe & Gray, 2007; Guided Search 5.0 -- Wolfe et al, 2015)。现有模型(见图 1)中自上而下影响因素主要由三个部分组成，包括模板引导(template guidance)、情景引导(episodic guidance)和语义引导(semantic guidance)。模板引导指的是搜索者对搜索目标物的了解和背景知识(Bahle, Matsukura, & Hollingworth, 2018; Duncan & Humphreys, 1989)。情景引导指的是在相似情境中目标物曾经在哪里出现过(Brooks, Rasmussen, & Hollingworth, 2010; Vo & Wolfe, 2012)。语义引导指的是在同类情境中目标物可能出现在哪里，受到背景信息、物体−场景关系和物体−物体关系的影响(Wolfe, Cain, & Aizenman, 2019; 见 Wu, Wick, & Pomplun, 2014 的综述)。至此，对影响真实环境中搜索的自上而下的因素分析已经比较完善，并形成一些统一的认识。

引导搜索模型中，自下而上的影响因素主要包括搜索对象的影像显著性(Koehler, Guo, Zhang, & Eckstein, 2014)，即搜索对象间影像差异性越大或者搜索目标的影像越突出，搜索越快越准。Koch 和 Ullman (1987)提出了显著性地图(saliency map)，即根据搜索对象多个特征的显著差异形成分布图，以此来预测观察者的搜索位置。该理论被许多实验研究结果支持(De Vries, Hooge, Wertheim, & Verstraten 2013; Kamkar, Moghaddam, & Lashgari, 2018)。然而，近年来有研究发现，影像显著性对注意分配的影响仅局限于实验室任务，不能泛化到真实环境中的视觉搜索。例如，在真实场景中搜索时，视觉刺激的显著性不能预测或解释观察者搜索时的眼动(Wu et al., 2014)。鉴于眼动注视点是体现注意分配的重要行为指标(Henderson & Hayes, 2017)，这意味着显著性地图模型不能解释真实搜索中的注意分配，也不能预测真实环境中的搜索行为(Foulsham & Underwood, 2009; Henderson, Malcolm, & Schandl, 2009)。

我们认为，通过实验室任务验证的显著性之所以无法解释真实环境中的搜索，是因为实验室任务中对于观察者和观察对象的设置太过简单和

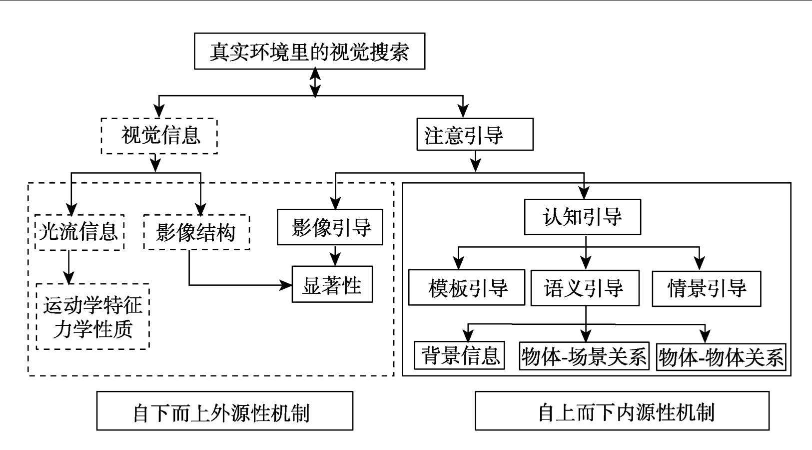


图 1 理论模型框架。我们认为视觉信息对真实环境里的视觉搜索起到自下而上的影响，尤其是光流以及光流和影像结构的交互。本研究补充了光流信息在视觉搜索中的自下而上的作用，但是运动产生的视觉、运动觉及本体觉信息可能以其它方式对视觉搜索造成自上而下的影响，有待后续研究。图中实线框为现有理论模型，主要基于 Wolfe 的引导搜索模型。虚线框内为本文提出理论构想的关键词。

理想化。首先，有研究者提出显著的观察对象必须要进入观察者的中央视野，其显著性才能起作用(Wolfe, 2003; Foulsham, Chapman, Nasiopoulos, & Kingstone, 2014)。在实验室电脑上进行的搜索任务，观察者大都端坐于电脑前，头部固定，静态的观察对象落在中央视野，这显然不是真实世界搜索时的情况。举例来说，Foulsham 等人(2014)设计了一个真实的搜索任务，要求参与者从实验室穿过几条走廊进入收发室，收发室里有一整面墙都是同样形状大小信箱格子(共 120 个)，被试需要找到一个目标信箱。在一半试次中，研究者将目标信箱涂成荧光粉色，希望明显的颜色特征可以突出目标，激发前注意搜索。但是实验结果显示，无论目标信箱有没有被涂成粉色，搜索反应时都是一样的，目标物的显著性对反应时没有影响。研究者认为，在真实场景中，观察者相对于环境来讲非常小，所以观察者需要先移动身体和头，对环境进行扫描，然后再用眼睛进行搜索。眼睛的搜索是嵌套在身体搜索中的二级搜索。当头刚好对着目标方向再用眼睛进行搜索时，显著性才起到作用。由于第一阶段的身体搜索远比眼睛搜索耗时长(在上述实验中，身体搜索用了 26 秒，而眼睛搜索用了 4 秒)，显著性对真实环境中搜索的帮助便无法体现出来了。第二，显著性不是一成不变的，而是动态的(dynamic)、情境的(situated)，会随着时间空间变化而变化。现有理论认为搜索对象在颜色、形状、大小、运动等几个维度上的差异影响搜索效率，差异越大目标越显著。但是在现实环境里，观察时间、观察者的姿态或物体的移动可能改变光影、观察角度、观察距离、观察对象间的遮挡关系等，从而改变搜索对象的颜色、形状、大小等影像信息。在这种情况下，搜索对象的显著程度几乎无法定义或量化。另外，真实搜索中也会存在多个维度一起变化的情况，比如一个物体的颜色突出，但是另一个物体是运动的，那么颜色和运动哪个更为显著就难有定论。

总结现有视觉搜索研究进展及对今后搜索研究的建议，我们认为要想解释真实环境中的视觉搜索，必须将搜索还原到最真实的环境(Kingston et al, 2003; Kingston, Smilek, & Eastwood, 2008)，加入环境的变化和观察对象的运动、考虑观察者的主动运动(Tatler, Hayhoe, Land, & Ballard, 2011)、重视观察者和环境的融合(Nakayama & Martini, 2011)。这些都是 Gibson 生态知觉理论的核心思想(Gibson, 1958; Gibson, 1979/1986)。所以将生态知觉理论引入视觉搜索显得自然且必须。在生态知觉理论体系下，运动产生了动态光流信息，这种信息可以标示环境中物体的结构和关系。动态光流信息和三维物体的静态影像结构(影像显著性)两者结合在一起形成了视觉信息。动、静态视觉信息作为一个整体取代原来基于平面图片(或场景)的影像显著性在引导搜索模型中的位置，构成影响视觉搜索的自下而上的因素。

2.2 生态知觉理论和视觉信息

在视觉搜索领域，复原真实环境的手法通常

是从场景照片里面搜索。但是 Gibson 明确提出，真实场景和场景图片是不同的：“看着尼亚加拉瀑布的感受和看着尼亚加拉瀑布照片的感受是不一样的(Gibson, 1979)”；同理，看着厨房的图片搜索和真正在厨房里搜索也是不一样的。差异在于在真实环境中，观察者或观察对象是可以运动的。运动使观察对象的影像不断变形、相互遮掩，使单纯基于影像匹配的搜索不可行。但是运动产生了另一种信息：光流，这是一种动态的信息，可以标示环境中物体的结构和关系。

生态光学理论认为一切视觉任务都依赖光学信息。Gibson (1966)提出光投入环境，被环境中的表面或物体反射，形成环境光(Ambient light)。环境光携带关于整个环境的信息。如，瓷砖、大理石、金属表面反射出的环境光不一样，所以通过察觉环境光，人可以知道哪个是厨房墙面，哪个是台面，哪个是洗菜盆。

环境光汇聚到一个观测点，形成一组光阵。对于某一观测点，构成静态光阵的各部分表面有着不同的视立体角(visual solid angle)，这些视立体角与环境中的物体表面的布局结构一一对应，形成静态影像结构信息。静态影像结构信息包括边界(edge)、光影(shading)、颜色或强度对比(contrast of color or intensity)等。这种信息是持久的，只要物体存在，影像结构信息就存在。

当观察者行进或环境中物体发生运动时，光阵中的各视立体角也随之发生变化，它们或新增、或消失、或放大、或缩小。光阵连续变化形成光流信息。光流状态与观察者在环境中相对运动速度、运动方向以及观察者与运动物体的距离一一对应，如距离观察者越远的物体光流速度越慢，在观察者正前方的物体比在她视野边缘的物体光流速度快。光流由运动产生，与运动模式一一对应；观察者通过察觉光流的状态、方向、速度和不动点的位置，知觉自身或环境物体的运动模式(图 2)。

环境中各物体表面对应某一个观测点并形成唯一的光阵，而观测点或环境物体的运动方式形成唯一的光流。这样的一一对应关系是由自然法则所决定的。环境中某一表面投射到某一观测点的影像结构信息由几何规律所约束，不是随机的；运动产生的连续光流由动力学规律和运动学规律约束，也不是随机的。这样的规律性，使观察者可以通过静态和动态信息准确知觉环境的结构和性质。

所以，要在实验室中复原真实环境里的搜索，仅使用真实图片或者虚拟现实显示是不够的，关键是要提供影像结构和光流，这样才能构建更具生态效度的实验场景，设置更接近真实环境的搜索条件，以探究真实环境里的视觉搜索。

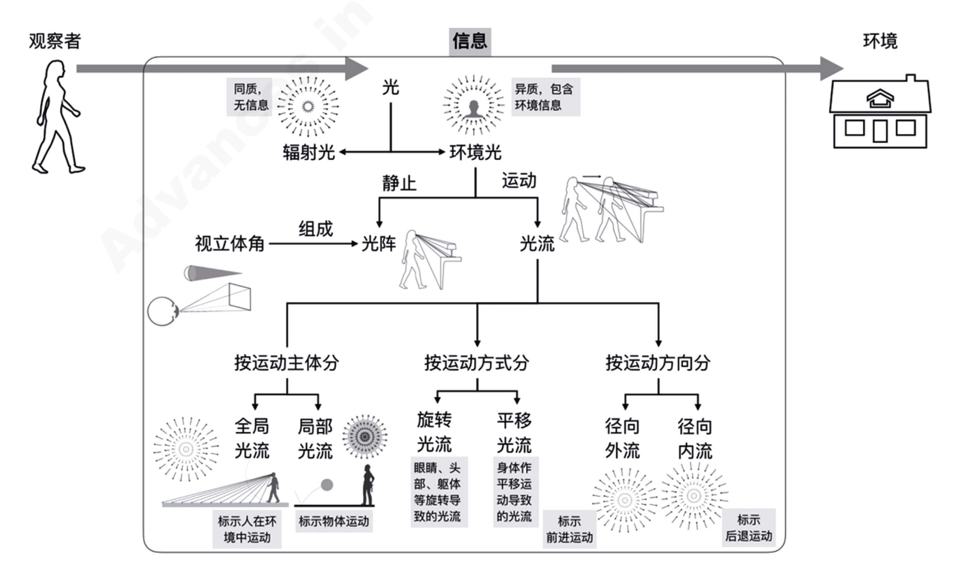


图 2　生态光学理论总结。Gibson 认为观察者利用环境光里的视觉信息完成知觉任务。环境光里包括静态影像结构信息和动态光流信息。

2.3 动、静态视觉信息在知觉场景、物体结构和事件中的作用

在自然的观察环境里，动、静态视觉信息同时存在。光流和影像结构信息结合可以帮助准确稳定地知觉场景、物体结构和事件。首先，研究发现观察者可以利用自身运动生成的光流识别模糊场景，光流强度和场景识别的表现呈正相关(Wu, Wang, & Pan, 2019)。

第二，观察者可以利用动态视觉信息准确知觉物体的三维结构，这个机制叫做运动恢复结构(structure-from-motion; Domini, Vuong, & Caudek, 2002; Todd, Tittle, & Norman, 1995)。如，Lind 和同事(Lee, Lind, Bingham, & Bingham, 2012)发现当将不同宽度−深度比的圆柱体摆在观察者面前时，观察者从约 45°俯视角观察，如果观察目标和观察者都处于静止状态，仅凭影像结构无法知觉物体的三维结构。但是只要观察者和观察目标之间存在连续的 45°以上的视角变化(观察者或者观察目标旋转 45°以上)，则观察者可以准确知觉物体三维结构。

第三，光流和影像结构可以帮助事件识别(Pan, Bingham, & Bingham, 2013; Pan et al, 2017)。事件指运动中的物体。而生物运动是事件的一种，也是其中最被深入研究的一种。研究发现观察者可以通过光点运动产生的视觉信息识别多种运动、动作者的特征、和其它非生物运动的事件(如，一个滚动的球，波动的水面等。Bingham, Rosenblum, & Schmidt, 1995)。此外，少量研究在生物运动范式中加入简单的影像信息(如将光点连线，或者加上轮廓线)，并使用贝叶斯模型(“理想观察者模型”, ideal observer model)分析了实验刺激里的信息含量，发现改变视觉刺激的信息量会影响识别、分辨生物运动的效率(Gold, Tadin, Cook, & blake, 2008; Lu, Tjan, & Liu, 2017)。

学者们对动、静态视觉信息在事件识别中所起到的作用进行了研究，并提出了“力约束运动理论” (kinematics-specified-by-dynamics theory, Runeson & Frykholm, 1983)。研究者们认为，因为每项运动背后的物理动力(dynamics)不同(如，形成跑、跳、走等不同运动的力是完全不同的)，每个运动者的身体具有不同的物理属性(质量、肢体长度、关节灵活性、肌肉强度等)，所以不同运动者做出的不同运动就具有独特的、固定的运动学(kinematics)特性。当观察者只接收到动态视觉信息时，根据观察到的运动特征，可以知觉生物运动及运动者本身的性质，从而知觉具体事件。Bingham 和同事后续提出标示事件动态信息的视觉信息是轨迹形态，即运动物体位置和速度的关系 $\left(\frac{\dot{x}}{x}\right)$, (trajectory form; Bingham et al., 1995; Bingham et al., 1995; Muchisky & Bingham, 2002; Wickelgren & Bingham, 2004, 2008b)。轨迹形态受力的影响。每种动力会产生唯一的轨迹形态，所以可以用来标示事件。研究发现，人们对这种动态信息非常敏感，可以通过该信息分辨非常类似的事件，如手晃动控制的钟摆和自由摆动的钟摆(Muchisky & Bingham, 2002)。更重要的是，轨迹形态信息不受观察视角影响，即便是从不熟悉的观察角度也可以识别事件(Wickelgren & Bingham 2004, 2008)。

综上所述，观察者的运动可以产生光流；通过运动恢复物体的结构，识别静止的三维物体；观察者注意物体的运动状态，可以识别事件及所涉及的物体属性。那么动、静态视觉信息在知觉活动中的作用是否可以迁移到视觉搜索任务中呢？答案是肯定的。前文已论述过，真实环境里的搜索和基于照片的搜索最大的区别就是运动。在真实环境里，观察者和观察对象之间的相对运动会改变静态影像信息及影像的显著性，所以有效的搜索机制必须可以适应或抵抗由运动引起的搜索对象外形的变化(Seidl-Rathkopf, Turk-Browne, & Kastner, 2015)。因而，独立于影像的、可以抵抗视角变化的动态视觉信息(如轨迹形态)很有可能就是人们在真实搜索中需要的信息。

2.4 动、静态视觉信息在视觉搜索任务中的作用

只包含运动信息的生物运动范式已经被应用于注意研究领域(Ding, Yin, Shui, Zhou, & Shen, 2017; Mayer, Vuong, & Thornton, 2015)，但使用生物运动范式研究视觉搜索的并不多。这些研究发现，观察者仅通过光点显示就可以从随机运动的光点中找出光点行走者(Hirai & Hiraki, 2006)，从正立的行走者里找出倒立的行走者(Wang, Zhang He, & Jiang, 2010)，找出行走方向不同的人(Cavanagh, Labianca, & Thornton, 2001)，和区分不同运动(Van Boxtel & Lu, 2011)。

真实环境中的视觉搜索通常同时存在动、静

态两种视觉信息。然而, 早期在引导搜索模型框架下的实验任务多采用二维静态图形或符号作为搜索对象, 观察者多数不能随意移动身体进行搜索。然而在真实环境里的视觉搜索, 对象可以是静止或移动的三维物体, 观察者也可以是静止或者运动的。例如在上述 Foulsham 等人(2014)的研究中, 虽然被试走进收发室, 寻找目标邮箱, 但是实验中也仅仅通过改变目标邮箱的显著性(加粉色边框)来操纵自下而上的信息, 却没有考虑移动过程中动态信息和静态信息的交互作用。近年来, 一些研究者采用不同方法探究被试走动对搜索的影响(Smith, Hood, & Gilchrist, 2008, 2010)。例如, Ruddle 和 Lessels (2006)借助仿真技术设计了一个被试在虚拟现实场景中搜索目标物的任务, 被试需要到 16 个不同位置寻找 8 个目标物。第一组被试只能坐在屏幕前(身体被固定)通过移动鼠标来模拟场景中的转向和前进运动; 第二组被试允许在固定位置上转动身体, 并借助立体显示头盔(Stereo HMD)实现在场景中的转向, 但仍需移动鼠标才能前进到不同位置; 不同于前两组, 第三组被试可以在真实环境中走到任意位置进行搜索。实验结果表明, 身体被固定的被试在搜索效率上均差于其他两组被试, 而允许随意走动的被试, 其搜索效率最高。也就是说, 身体运动所产生的动态信息与搜索对象的静态信息相结合时可能有利于提高视觉搜索效率。

前人提出的“具身记忆模型” (Pan et al., 2013)认为, 当动、静态视觉信息同时标示一个事件时, 光流具有空间方面的准确性, 可以校准影像结构, 帮助观察者准确识别物体和环境的三维关系; 影像结构具有时间方面的稳定性, 可以在运动停止、光流消失后形成具身记忆, 使观察者持续知觉三维结构。我们发现影像信息和光流信息的结合可以使观察者准确找出被隐藏或被伪装的目标物。在 Pan 等人(2013, 2017)的研究中, 多个目标物逐渐被干扰物遮挡, 被试可以利用两种视觉信息, 在遮挡的过程中和被完全遮挡后, 准确找出目标物。在 Pan, Bingham, Chen 和 Bingham (2017)的研究中, 当目标物和干扰物外形完全一样, 但空间位置不同时, 被试可以利用两种视觉信息准确稳定地找出目标物。按照 Nakayama 和 Martini (2011)的对识别和搜索任务的界定(识别是通过多种特征认出一个目标物, 搜索是通过少数特征找出多个目标物), 上述两个任务实际上更偏向于视觉搜索。

综上, 现有研究已经表明, 观察者可以利用动态视觉信息搜索事件(生物运动及非生物运动), 可以利用静态视觉信息搜索物体。于是, 我们在现有引导搜索模型的基础上, 加入视觉信息变量, 提出光流和影像结构是影响搜索的自下而上的重要因素。观察者利用搜索对象之间影像的差异区分不同个体, 差异越大的物体越显著, 容易被搜索; 观察者可以利用光流信息知觉搜索对象的运动特征及背后的力学属性, 从而区分不同的搜索对象(图 1)。当然, 观察者自身的运动除了可以在视觉层面生成光流信息, 也会产生运动觉、本体觉等信息。观察者在运动中进行视觉搜索时, 运动提供了更多的自下而上的视觉信息, 帮助视觉搜索; 但不能排除运动相关信息对高级认知过程的影响, 如工作记忆、注意等, 且这种影响可能是抑制性的(Mayer, Riddell, & Lappe, 2019)。

3　研究构想

3.1　科学问题

本研究拟解决的第一个科学问题是：影响真实环境中搜索三维物体和事件的自下而上的因素是什么。现有理论将影响真实环境中搜索行为的因素分为自下而上和自上而下两类, 其中自下而上的原因被归结为搜索对象影像的显著性。我们认为这是不充分的。我们提出对搜索产生自下而上影响的是视觉信息, 包括静态影像结构和动态光流信息, 且两种信息之间存在交互作用。我们通过研究一和研究二分别研究影像结构信息和光流信息在搜索静止的三维物体和运动的事件时的作用, 从而回答动、静态视觉信息如何被整合利用以完成真实环境中的视觉搜索这一科学问题。

本研究拟解决的第二个科学问题是：传统的视觉搜索理论能否泛化并预测真实环境中的视觉搜索行为。传统的视觉搜索研究多使用二维图像作为搜索对象, 经过几十年的探索得出许多理论, 这些理论可以解释基于二维图像的搜索行为, 如放射科医生从 X 光片中识别异常组织。但是真实环境中的搜索任务更加复杂：搜索目标、干扰子是三维的, 观察者和观察对象是可以运动的, 搜索视角会变化, 背景环境繁杂等等。所以在平面图像上的视觉搜索和真实环境中的视觉搜索是否

存在相似的行为规律？我们通过直接比较两种搜索表现(研究一)和比较搜索训练的效果(如果通过训练平面上的搜索可以提高真实环境中的搜索表现，则两种搜索本质上相通；研究三)两种方法回答这个问题。

3.2 研究方案

在本项目中，我们通过三个子研究(技术路线总图见图 3)来探究利用动、静态视觉信息的视觉搜索过程和机制，以及提高视觉搜索效率的干预方式。研究中主要使用心理物理法、贝叶斯估计等方法分析刺激中的信息量、搜索效率及二者之间的关系。通过三个子研究，系统梳理影像结构和光流信息在搜索中的作用，验证视觉信息对搜索产生的自下而上的影响，完善基于视觉信息和注意引导的搜索模型。再将理论应用到对人的视觉搜索训练，和对智能搜索的设计中去，从而帮助人和机器更好地完成搜索任务。

3.2.1 研究一：利用动、静态视觉信息对静止目标物体的搜索

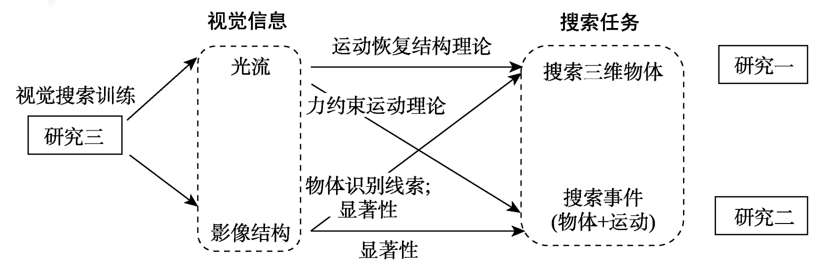
研究一将通过 3 个实验来探索当搜索目标静止时，静止或运动的观察者如何进行视觉搜索，找到目标物体。分别对应三个问题：1)运动恢复结构能否帮助搜索三维物体；2)视觉搜索效率是否受视角转变(perspective change)的影响；3)观察者运动时，视角发生连续变化，导致视网膜上的像连续变化，两种视觉信息能否解决这种变化对搜索的影响。

在实验 1.1 中，我们在电脑屏幕上分别以正投影(观察角度为 0)和透视投影(45°俯视角)的方式呈现刺激，比较搜索表现以得出视角转变是否对视觉搜索产生影响。实验 1.2 中，我们将真实物体排列在桌面上，被试坐在桌边，以 45°俯视角观察搜索序列，在搜索对象不动、搜索对象被动旋转、及被试主动旋转搜索对象的情况下，找出目标物体。对比实验 1.2 和实验 1.1，我们可以得知真实环境中的搜索和电脑上模拟的真实场景中的搜索是否相同，视角变化及运动恢复结构能否促进搜索。实验 1.3 中，我们在虚拟现实环境中搭建更逼真更复杂的搜索场景，允许观察者自由移动观察。使用与实验 1.1、1.2 相似的条件，比较三个实验的结果，以验证实验室研究的效度，探索影像结构和光流信息结合对真实环境中物体搜索的影响。

3.2.2 研究二：利用动、静态视觉信息对运动事件的搜索

研究二的主要问题是：当搜索对象是运动中的人时(运动的人是事件)，观察者如何利用视觉信息找到某一个人。许多利用生物运动范式的研究指出，人类观察者对人的运动十分敏感，可以仅通过运动信息(而不需要影像信息)识别人的动作，对动作进行分类，或分辨运动者的性别、体型、情绪等。那么，我们可以仅通过运动信息从一群运动的人中找出某一个运动的个体吗？此外，在传统的生物运动范式中，一个运动的人被简化成一组协动的光点，并从矢状面(sagittal view)以正投影方式呈现，且多数情况下，光点组之间是独立的、没有重叠或穿插(如一个或几个独立的光点运动者在电脑屏幕上向左或向右走)。但是在真实环境中，多人运动不只停留在额平行平面(frontoparallel plane)上，也有纵深运动，且会有很多重叠遮掩。而观察者的视线和搜索对象所在平面也可能不垂直，如站在高处看或者监控拍到





研究思路: 控制视觉信息, 比较搜索效率; 针对不同信息的训练能否提高真实搜索的表现

实验手段: 心理物理法、光流量化算法、贝叶斯模型、眼动、虚拟现实、裸眼3D

图 3 本项目的技术路线总图。其中分三个研究，分别对应静态三维物体搜索、事件搜索以及视觉搜索训练。

的视频通常存在俯视角。所以，当存在纵深维度运动和运动者相互遮挡时，以及当观察者和搜索对象之间存在视角变化时，观察者能否依赖运动信息搜索到某一个运动的目标？第三，在真实环境中，搜索对象不仅有运动信息还有影像信息。前人研究发现，生物运动的呈现方式(光点、连线、轮廓线或剪影)对动作识别效率有很大影响，加入一些影像信息(如将光点之间连线再呈现)会使识别效率更高(Lu et al., 2017)。同时，影像结构还能影响目标物和干扰子之间、干扰子和干扰子之间的相似程度以及搜索对象的显著性，从而改变搜索任务难度。综合这两方面的原因，在搜索任务中，加入影像信息对运动目标的搜索会产生怎样的影响呢？最后，在真实环境中，往往存在不止一种运动，如在街上有人行走，也有汽车行驶。根据“力约束运动”理论，人行走和汽车行驶的运动本质上是不同的，具有完全不同的力学属性，可以很容易从运动信息中区分两种运动。在信息层面，加入另一种完全不同的运动不会影响搜索。但是运动是一种显著的线索，运动的干扰子可以瞬间抓住观察者注意力，影响搜索效率。所以，目标显著性和运动学特征两种自下而上的影响对事件搜索产生怎样的交互影响呢？

我们将通过 4 个实验回答上述问题。实验 2.1 结合视觉搜索和生物运动的实验范式，将运动的光点组相互独立地从矢状面方向以正投影方式呈现在屏幕上，要求被试从几组运动的光点中找出目标运动者。实验 2.2 中，搜索对象在空间中穿插走动时，光点间会出现重叠、穿插及非刚性运动(non-rigid motion)。实验材料会从 0°和 45°两个俯视角制作，模拟沉浸在人群中的观察者的搜索视角和监控录像中的搜索视角两种情况。被试通过光点组的运动找出目标事件。在实验 2.3 中，我们用虚拟现实搭建一个在繁忙场所找人的情景，赋予搜索对象影像结构信息，操控影像显著性(如改变被搜索人群衣服的颜色或统一性)，比较观察者静止观察时和运动观察时的事件搜索表现。在实验 2.3 的基础上，实验 2.4 中包含静态和动态干扰子，研究干扰子的影像显著性和干扰子的运动特性对真实环境下事件搜索的交互影响。我们加入运动的干扰子(如在街上找某个行人的搜索任务中加入行驶的车辆)，以探究无关运动信息对搜索的影响；然后改变静态干扰子的显著度(如加入闪烁的路边招牌)，以探究影像结构信息显著性对搜索的影响。

3.2.3　研究三：运用仿真手段对视觉搜索的训练

研究一、二从理论层面探究动、静态视觉信息如何被整合利用从而完成在真实环境中搜索静止的物体或动态的事件，并梳理总结各种因素对搜索效率的影响。在此理论基础之上，研究三旨在找出能够有效提高视觉搜索效率的训练方法。搜索训练的目标是提高在复杂搜索任务中的正确率和搜索效率，搜索任务包括静止或运动的观察者搜索目标物体或事件。训练分为 4 个阶段：前测–训练–后测–保持。前测指未经训练的观察者在完成任务时的基线水平。后测指经过训练后，达到的水平。保持指在训练后的一段时间里，较高的搜索效率能否持续。训练是最重要的一个阶段，设计训练方案的一个重要理念是如何通过练习简单任务提高在复杂任务中的表现。训练需根据理论研究得出的影响搜索的因素，通过简单的、可操控的、有针对性的训练任务，来提高复杂搜索任务的表现。同时，训练效果也可以进一步证明理论研究中得到的因素是否确实对搜索有影响。

研究三通过 3 个实验找出最有效的训练方法。每个实验的前侧、后测、保持阶段的任务为：静止或运动的观察者搜索静止的目标和静止的观察者搜索运动的目标(分别对应研究一、二)，而训练阶段依照训练内容的复杂程度分为三个实验。实验 3.1：训练阶段使用虚拟现实搭建的、跟其它三阶段一样的仿真环境。实验 3.2：训练阶段使用抽象的三维搜索对象，但仍然有光流和影像结构信息，如研究一、二中使用的搜索乐高积木或搜索光点行走者任务。实验 3.3：训练阶段使用传统的平面视觉搜索范式(如在彩色形状中找红色圆圈之类)，视觉信息不同，但是搜索同样需要注意分配和控制。通过实验 3.1，我们可以知道能否通过训练提高真实环境中的视觉搜索。通过实验 3.2，我们可以知道针对视觉信息的训练，能否提高视觉搜索表现。通过实验 3.3，我们可以知道针对搜索过程中注意的训练，能否提高搜索表现。从实验 3.1 到实验 3.3，知觉信息逐步减少但是搜索任务原理上一致，所以综合比较三种训练的效果，我们可以间接知悉影响搜索的视觉信息因素和注意因素之间的关系，从而间接验证我们提出的基于视觉信息和注意机制的搜索理论。

4 理论建构

成功的视觉搜索是人类生存繁衍的必要技能。大量视觉搜索研究基于引导搜索模型开展。该模型认为：在搜索过程中，对注意的引导可分为自上而下和自下而上两个部分。自上而下的影响因素包括模板引导、情景引导和语义引导三个部分。对于这三个部分的研究已经比较完善并形成一些统一的认识。自下而上部分则简单归为“影像显著性”。然而，近年来有研究发现，影像显著性对注意分配的影响局限于实验室任务，不能泛化到真实环境中的视觉搜索(Henderson & Hayes, 2017; Wu, Wick, & Pomplun, 2014)。影响视觉搜索的自下而上因素是什么则成了关键问题。另外，传统视觉搜索研究多数针对静止观察者和静止的搜索对象，且搜索对象以平面形式呈现。然而这只是视觉搜索的一种情况。真实环境中的视觉搜索，搜索对象可以是静止或移动的三维物体，观察者也可以是静止或者运动的。传统的实验室研究结果能否泛化并预测真实环境中的视觉搜索也有待解答。

为了解决上述两个问题，本团队特地设计了三个研究进行回答。我们将结合生态知觉理论(Gibson, 1966, 1979)，引入动、静态视觉信息来完善自下而上的影响因素并提出理论模型(见图 1)。同时，希望在视觉搜索研究领域取得以下进展：

首先，研究一包括静止观察者搜索静止三维物体(可利用信息为影像结构)，和运动观察者搜索静止三维物体(可利用信息为影像结构和全局光流)两种情况。根据研究一的行为数据，我们将探索真实环境下动、静态视觉信息对搜索三维物体的影响，了解搜索三维物体的行为规律。再根据 Pan, Bingham 和 Bingham (2013, 2017)的“具身记忆模型”理论，比较基于平面结构的搜索和真实环境下三维物体的搜索的异同，以验证传统平面搜索研究得出的理论是否适用于搜索三维物体。该研究的结果一方面可以说明光流对视觉搜索所起的作用，弥补之前理论模型中的缺失。另一方面也能说明，影像结构可以保存光流标示的物体或事件，使搜索具备持久性。

其次，研究二将设计两个任务：静止观察者搜索运动三维物体，即事件(可利用信息为影像结构和局部光流)；运动观察者搜索事件(可利用信息为影像结构和全局及局部光流)，结合生物运动范式，验证真实环境下动、静态视觉信息在事件搜索中的作用。我们验证光流对运动特征及其背后的力学性质的标示能否是观察者区分事件，并通过与影像信息的结合，实现准确、持久的视觉搜索；此外，在加入不同扰动(如，视角变化，遮掩等)之后，可以探究光流信息对事件标示的抗变换性。最后，我们将在虚拟仿真环境下提供静态影像结构和光流信息，以获知在复杂场景中的搜索过程和规律。

最后，在前两个研究的基础上，我们将在研究三中找出能够有效提高视觉搜索效率的训练方法。其中，第一个实验验证在虚拟现实场景中重复练习视觉搜索能否提高搜索效率；第二个实验强调对影像结构和光流信息的提取及整合能力的训练，通过练习搜索运动的三维搜索对象，检验训练能否提高在虚拟仿真场景中复杂搜索任务的表现。第三个实验则采用的传统视觉搜索范式，通过练习对平面图形或符号的搜索，检验训练能否提高在虚拟仿真场景中复杂搜索任务的表现。如果实验二和实验三的训练效果存在差异，则间接说明传统的视觉搜索范式不能完全代表三维环境中的视觉搜索行为。

综上所述，本研究将使用多种研究技术和手段，系统梳理影像结构和光流信息在搜索中的作用，探究动、静态视觉信息对搜索产生的自下而上的影响，完善基于视觉信息和注意指引的搜索模型。在本研究基础上，后续研究可深入发掘观察者自身运动产生的视觉、本体觉、运动觉等信息对搜索任务造成的自上而下的影响，进而完善具有高生态效度的视觉搜索理论体系。例如，寻觅(foraging)任务被认为是近似于视觉搜索的一种自然任务。在寻觅任务中，观察者可以移动身体反复查看并找出目标(Ehinger & Wolfe, 2016; Wolfe, Cain, Ehinger, & Drew, 2015)，此时运动相关信息与高级认知功能(注意、记忆等)交互就显得格外重要了。另一方面，生态知觉理论并没有区分生物运动和非生物运动，它们都是事件，都可以通过捕捉动态轨迹形态信息识别。但是我们目前规划的视觉搜索实验仅使用了生物运动作为搜索对象。后续研究可将搜索对象扩展到其他类型的事件，从而验证通过动态视觉信息搜索事件这一理论的普遍性。

参考文献

Bahle, B., Matsukura, M., & Hollingworth, A. (2018). Contrasting gist-based and template-based guidance during real-world visual search. *Journal of Experimental Psychology: Human Perception and Performance, 44*(3), 367–386.

Bingham, G. P., Schmidt, R. C., & Rosenblum, L. D. (1995). Dynamics and the orientation of kinematic forms in visual event recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*(6), 1473-1493.

Broadbent, D. E. (1991). A word before leaving. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV* (pp. 863–879). Cambridge, MA: Bradford Books/MIT Press.

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1406–1418.

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition, 80*(1-2), 47–60.

De Vries, J. P., Hooge, I. T. C., Wertheim, A. H., & Verstraten, F. A. J. (2013). Background, an important factor in visual search. *Vision Research, 86*, 128–138.

Ding, X., Yin, J., Shui, R., Zhou, J., & Shen, M. (2017). Backward-walking biological motion orients attention to moving away instead of moving toward. *Psychonomic Bulletin & Review, 24*(2), 447–452.

Domini, F., Vuong, Q. C., & Caudek, C. (2002). Temporal integration in structure from motion. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 816-838.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2016). Searching while loaded: Visual working memory does not interfere with hybrid search efficiency but hybrid search uses working memory capacity. *Psychonomic Bulletin & Review, 23*(1), 201–212.

Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of Vision, 17*(11), 5.

Duncan, J. S., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–458.

Ehinger, K. A., & Wolfe, J. M. (2016). When is it time to move to the next map? Optimal foraging in guided visual search. *Attention, Perception, & Psychophysics, 78*(7), 2135–2151.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*(4), 249–263.

Foulsham, T., Chapman, C. S., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real-world environment. *Canadian Journal of Experimental Psychology, 68*(1), 8–19.

Foulsham, T., & Underwood, G. (2009). Does conspicuity enhance distraction? Saliency and eye landing position when searching for objects. *The Quarterly Journal of Experimental Psychology*, *62*(6), 1088–1098.

Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology, 49*(3), 182–194.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc. (Original work published 1979).

Gold, J. M., Tadin, D., Cook, S. C., & Blake, R. (2008). The efficiency of biological motion perception. *Attention Perception & Psychophysics, 70*(1), 88–95.

Henderson, J. M., & Hayes, T. R. (2017). Meaning-based guidance of attention in scenes as revealed by meaning maps. *Nature Human Behaviour, 1*(10), 743–747.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review, 16*(5), 850–856.

Hickey, C., Chelazzi, L., Theeuwes, J., & Geng, J. J. (2014). Reward-priming of location in visual search. *PLoS ONE, 9*(7), e103372.

Hirai, M., & Hiraki, K. (2006). Visual search for biological motion: An event-related potential study. *Neuroscience Letters, 403*(3), 299–304.

Kamkar, S., Moghaddam, H. A., & Lashgari, R. (2018). Early visual processing of feature saliency tasks: A review of psychophysical experiments. *Frontiers in Systems Neuroscience, 12*, 54.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology, 99*, 317–340.

Kingstone, A., Smilek, D., Ristic, J., Friesen, C. K., & Eastwood, J. D. (2003). Attention, researchers! It is time to take a look at the real world. *Current Directions in Psychological Science, 12*(5), 176 –180.

Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology, 4*(2), 115–141.

Koehler, K., Guo, F., Zhang, S., & Eckstein, M. P. (2014). What do saliency models predict. *Journal of Vision, 14*(3), 14–14.

Kristjánsson, Á., Jóhannesson, Ó. I., & Thornton, I. M. (2014). Common attentional constraints in visual foraging. *Plos One, 9*(6), e100752.

Lee, Y. L., Lind, M., Bingham, N. & Bingham, G. P. (2012). Object recognition using metric shape. *Vision Research, 69*, 23–31.

Lu, H., Tjan, B. S., & Liu, Z. (2017). Human efficiency in detecting and discriminating biological motion. *Journal of Vision, 17*(6), 4–4.

Mayer, K. M., Riddell, H., & Lappe, M. (2019). Concurrent processing of optic flow and biological motion. *Journal of Experimental Psychology: General*, *148*(11), 1938–1952.

Mayer, K. M., Vuong, Q. C., & Thornton, I. M. (2015). Do People "Pop Out"?. *PLOS ONE, 10*(10), e0139618.

Muchisky, M. M. & Bingham, G. P. (2002). Trajectory forms as a source of information about events. *Attention Perception & Psychophysics, 64*(1), 15–31.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research, 51*(13), 1526–1537.

Ort, E., Fahrenfort, J. J., & Olivers, C. N. L. (2017). Lack of free choice reveals the cost of having to search for more than one object. *Psychological Science, 28*(8), 1137−1147.

Pan, J. S., Bingham, N., & Bingham, G. P. (2013). Embodied memory: Effective and stable perception by combining optic flow and image structure. *Journal of Experimental Psychology: Human Perception and Performance, 39*(6), 1638–1651.

Pan, J. S., Bingham, N., & Bingham, G. P. (2017). Embodied memory allows accurate and stable perception of hidden objects despite orientation change. *Journal of Experimental Psychology: Human Perception and Performance, 43*(7), 1343–1358.

Pan, J. S., Bingham, N., Chen, C., & Bingham, G. P. (2017). Breaking camouflage and detecting targets require optic flow and image structure information. *Applied Optics, 56*(22), 6410–6418.

Pan, J. S., Li, J., Chen, Z., Mangiaracina, E. A., Connell, C. S., Wu, H., ... Hassan, S. E. (2017). Motion-generated optical information allows event perception despite blurry vision in AMD and amblyopic patients. *Journal of Vision, 17*(12), 13–13.

Ruddle, R. A., & Lessels, S. (2006). For efficient navigational search, humans require full physical movement, but not a rich visual scene. *Psychological Science, 17*(6), 460–465.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General, 112*(4), 585–615.

Seidl-Rathkopf, K. N., Turk-Browne, N. B., & Kastner, S. (2015). Automatic guidance of attention during real-world visual search. *Attention Perception & Psychophysics, 77*(6), 1881–1895.

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2008). Visual search and foraging compared in a large-scale search task. *Cognitive Processing*, *9*(2), 121–126.

Smith, A. D., Hood, B. M., & Gilchrist, I. D. (2010). Probabilistic cuing in large-scale environmental search. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 36(3), 605−618.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5), 5–5.

Theeuwes, J., Kramer, A. F., & Belopolsky, A. V. (2004). Attentional set interacts with perceptual load in visual search. *Psychonomic Bulletin & Review, 11*(4), 697–702.

Todd, J. T., Tittle, J. S., & Norman, J. F. (1995). Distortions of three-dimensional space in the perceptual analysis of motion and stereo. *Perception, 24*(1), 75–86.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766−786.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology, 8*(2), 194−214.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95*(1), 15–48.

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. *Attention and Performance* VI, 333.

Van Boxtel, J. J., & Lu, H. (2011). Visual search by action category. *Journal of Vision, 11*(7), 19–19.

Vo, M. L., & Wolfe, J. M. (2012). When does repeated search in scenes involve memory? Looking at versus looking for objects in scenes. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 23–41.

Võ, M. L. H., & Wolfe, J. M. (2015). The role of memory for visual search in scenes. *Annals of the New York Academy of Sciences, 1339*(1), 72–81.

Wang, L., Zhang, K., He, S., & Jiang, Y. (2010). Searching for life motion signals: Visual search asymmetry in local but not global biological-motion processing. *Psychological Science, 21*(8), 1083–1089.

Wickelgren, E. A. & Bingham, G. P. (2004). Perspective distortion of trajectory forms and perceptual constancy in visual event identification. *Attention Perception & Psychophysics, 66*, 629–641.

Wickelgren, E. A., & Bingham, G.P. (2008). Trajectory forms as information for visual event recognition: 3D perspectives

on path shape and speed profile. *Attention Perception & Psychophysics, 70*(2), 266–278.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin &. Review, 1*, 202–238

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7*(2), 70–76.

Wolfe, J. M., Boettcher, S. E. P., Josephs, E. L., Cunningham, C. A., & Drew, T. (2015). You look familiar, but I don't care: Lure rejection in hybrid visual and memory search is not based on familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 41*(6), 1576–1587.

Wolfe, J. M., Cain, M. S., & Aizenman, A. M. (2019). Guidance and selection history in hybrid foraging visual search. *Attention, Perception, & Psychophysics, 81*(3), 637–653.

Wolfe, J. M., Cain, M. S., Ehinger, K. A., & Drew, T. (2015). Guided Search 5.0: Meeting the challenge of hybrid search and multiple-target foraging. *Journal of Vision, 15*(12), 1106–1106.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 419–433.

Wolfe, J. M., & Gancarz, G. (1997). Guided Search 3.0. *In Basic and clinical applications of vision science* (pp. 189–192). Dordrecht: Springer.

Wolfe, J. M., & Gray, W. (2007). Guided search 4.0. *Integrated models of cognitive systems*, 99–119.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour, 1*(3), 0058.

Wolfe, J. M., & van Wert, M. J. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology, 20*(2), 121–124.

Woodman, G. F., & Chun, M. M. (2006). The role of working memory and long-term memory in visual search. *Visual Cognition, 14*(4-8), 808–830.

Wu, C., Wick, F. A., & Pomplun, M. (2014). Guidance of visual attention by semantic information in real-world scenes. *Frontiers in Psychology*, *5*, 54–54.

Wu, H., Wang, X. M., & Pan, J. S. (2019). Perceiving blurry scenes with translational optic flow, rotational optic flow or combined optic flow. *Vision Research, 158*, 49–57.

Visual search in real world: The role of dynamic and static optical information

PAN Jing; ZHANG Huiyuan; CHEN Donghao; XU Hongge

(*Department of Psychology, Sun Yat-sen University, Guangzhou* 51006, *China*)

Abstract: Visual search is a ubiquitous task and a critical skill for men and animals. Existing studies on visual search mainly focus on attentional guidance and the top-down cognitive influences on search effectiveness. The bottom-up influence on visual search is, rather crudely, simplified as objects' image saliency. However, when searching in real world, where the observer and/or objects move, both static image information (the saliency of which has been considered in existing search models) and dynamic optic flow information are available. Optic flow is generated by the relative motions between an observer and world objects. So by detecting flow patterns, observers get to know the kinematic properties of events (which is defined as objects in motion) and hence perceive the physical properties of constituent objects, such as the mass, size and frictional coefficient etc.. These physical properties distinguish objects and allow the observer to search for a particular one. We integrate dynamical perceptual information (i.e. optic flow) into existing search models and in two studies, we test how combined dynamical and static perceptual information affect visual search for three-dimensional objects and for moving people, when the observer is stationary or moving. Furthermore, we attempt to develop a training protocol that improves search effectiveness in real world. Findings from this project will bring forth new theories for understanding visual search in real world, and have direct applications on personnel training and intelligent search designs.

Key words: visual search; optical flow; biological motion; ecological theory of perception